Chapter Review Exercises (page 278)

R4.1 (a) The population is all Ontario residents and the sample is the 61,239 people interviewed.
(b) Even though the sample size is very large, it is unlikely that the percentages in the entire population would be exactly the same as the percentages in the sample because of sampling variability. However, these two numbers should be fairly accurate reflections of the values for the whole population.

R4.2 Answers will vary. One possible answer is:
(a) Announce in daily bulletin that there is a survey concerning student parking available in the main office for students who want to respond. Since voluntary response surveys are generally responded to only by those who feel strongly about the issue, these results will likely be biased.
(b) Personally interview a group of students as they come in from the parking lot. Convenience samples are not generally representative of the population leading to biased results.

R4.3 (a) Number the players from 01 to 25 in alphabetical order: Agarwal = 01, Andrews = 02, …, Wilson = 25. Move from left to right reading pairs of digits until you find three different pairs between 01 and 25. The three players with the corresponding numbers would be chosen for the drug test.
(b) Using the numbers given, choose **17** (Musselman) 52 (skip) 17 (repeat) 80 (skip) **09** (Fuhrmann) 46 (skip) **23** (Smith).

R4.4 A stratified random sample would probably be best here because it is likely that the opinions of professors will vary based on which type of institution they are at. If that is the case, using a stratified random sample will provide more precise estimates than using a simple random sample or cluster sample. Furthermore, a simple random sample or a cluster sample might miss faculty from one particular type of institution, especially if there are not many faculty at that type of institution.

R4.5 (a) A potential source of bias related to the question wording is that people may not remember how many movies they watched in a movie theater in the past year. It might help the polling organization to shorten the amount of time that they ask about, perhaps 3 or 6 months.
(b) A sample that only uses residential phone numbers is likely to underrepresent younger adults who use only cell phones. If younger adults go to movies more often than older adults, the estimated mean will be too small.
(c) People who do not go to the movie theater very often might be more likely to respond to the poll because they are at home. Because the frequent movie goers will not be at home to respond, the estimated mean will be too small.

R4.6 (a) The data were collected after the anesthesia was administered. Hospital records were used to "observe" the death rates, rather than imposing different anesthetics on the subjects.
(b) Two variables are confounded when it is unclear which variable is causing a change in the response variable. One variable that might be confounded with choice of anesthetic is type of surgery. If anesthesia C is used more often with a type of surgery that has a higher death rate, we wouldn't know if the death rate was higher because of the anesthesia or the type of surgery.

R4.7 (a) The experimental units are the potatoes used in the experiment. The explanatory variables are storage method and time from slicing until cooking. The response variables are ratings of color and flavor. There are six treatments: (1) fresh picked and cooked immediately, (2) fresh picked and cooked after an hour, (3) stored at room temperature and cooked

immediately, (4) stored at room temperature and cooked after an hour, (5) stored in refrigerator and cooked immediately, (6) stored in refrigerator and cooked after an hour.
(b) Using 300 identical slips of paper, write "1" on 50 of them, write "2" on 50 of them, and so on. Put the papers in a hat and mix well. Then, select a potato and randomly select a slip from the hat to determine which treatment that potato will receive. Repeat this process for the remaining 299 potatoes, making sure not to replace the slips of paper into the hat.
(c) It would be best to use a randomized block design with regular potatoes in one block and sweet potatoes in the other block. Randomly assign the 6 treatments within each block as in part (b).

R4.8 (a) No. The 1000 students were not randomly selected from any larger population so we should not make an inference about a population.
(b) Yes. Because the students were randomly assigned to the three treatments, we can conclude that the reduction in cold symptoms was caused by the masks.

R4.9 (a) If all of the patients received the Saint-John's-wort, the researchers wouldn't know if any improvement was due to the Saint-John's-wort or to the expectations of the subjects (the placebo effect). By giving some patients a treatment that should have no effect at all, but that looks, tastes, and feels like the Saint-John's-wort, the researchers can account for the placebo effect by comparing the results for the two groups.
(b) The purpose of random assignment is to create two groups of subjects that are roughly equivalent at the beginning of the experiment.
(c) The subjects should not know which treatment they are getting so that the researchers can measure how much placebo effect there is. Also, the researchers should be unaware of which subjects received which treatment so they cannot consciously (or subconsciously) influence how the results are measured.
(d) In this context, "not statistically significant" means that the difference in improvement between the St.-John's-wort and placebo groups wasn't large enough to rule out random chance as the explanation. In other words, it is plausible that the difference was solely due to the variability caused by the random assignment to treatments.

R4.10 (a) Randomly assign 15 students to Group 1 (easy mazes) and the other 15 to Group 2 (hard mazes). To do this, use 30 identical slips of paper and write the name of each subject on a slip. Mix the slips in a hat and select 15 of them at random. These subjects will be assigned to Group 1. The remaining 15 will be assigned to Group 2. After the experiment, compare the time estimates of Group 1 with those of Group 2.
(b) Each student does the activity twice, once with the easy mazes, and once with the hard mazes. For each student, randomly determine which set of mazes is used first. To do this, flip a coin for each subject. If the coin lands on heads, the subject will do the easy mazes followed by the hard mazes. If the coin lands on tails, the subject will do the hard mazes followed by the easy mazes. After the experiment, compare each student's "easy" and "hard" time estimate.
(c) The matched pairs design would be more likely to detect a difference because it accounts for the variability between subjects.

R4.11 (a) This does not meet the requirements of informed consent because the subjects did not know the nature of the experiment before they agreed to participate.
(b) All individual data should be kept confidential and the experiment should go before an institutional review board before being implemented.